

Computing Quality of Experience of Video Streaming in Network with Long-Range-Dependent Traffic

Zakaria Ye*, Rachid EL-Azouzi*, Tania Jimenez* and Yuedong Xu†

*University of Avignon, 339 Chemin des Meinajaries, Avignon, France

†Department of Electronic Engineering, Fudan University, China

Email: {zakaria.ye, rachid.elazouzi, tania.jimenez}@univ-avignon.fr

ydxu@fudan.edu.cn

Abstract—We take an analytical approach to study the Quality of user Experience (QoE) for video streaming applications. Our propose is to characterize buffer starvations for streaming video with Long-Range-Dependent (LRD) input traffic. Specifically we develop a new analytical framework to investigate Quality of user Experience (QoE) for streaming by considering a Markov Modulated Fluid Model (MMFM) that accurately approximates the Long Range Dependence (LRD) nature of network traffic. We drive the close-form expressions for calculating the distribution of starvation as well as start-up delay using partial differential equations (PDEs) and solve them using the Laplace Transform. We illustrate the results with the cases of the two-state Markov Modulated Fluid Model that is commonly used in multimedia applications. We compare our analytical model with simulation results using ns-3 under various operating parameters. We further adopt the model to analyze the effect of bitrate switching on the starvation probability and start-up delay. Finally, we apply our analysis results to optimize the objective quality of experience (QoE) of media streaming realizing the tradeoff among different metrics incorporating user preferences on buffering ratio, startup delay and perceived quality.

I. INTRODUCTION

It has been observed that people addicted to watching streaming videos constitute more than half of the Internet traffic. With the introduction of smartphones, mobile networks are witnessing an exponential traffic growth every year. This leads to scenario where Internet and wireless networks are pushed to operate close to their performance limits, dictated by current architectural considerations. Though much effort has been expended and in turn significant progress has been made in recent years to increase the capacity of mobile networks, there is little progress on dealing with the user satisfaction, which is strongly related to the Quality of Experience (QoE). This is the big challenge that the operators face today because they have to look at both the server side and the client side to make a link between the quality of service (QoS) of the network and the client satisfaction which depends on the QoE. Empirical studies in [4], [7], [11], [12] has identified critical metrics that affect the QoE through the user engagement:

- Starvation probability. Denoting the probability that a streaming user sees frozen images.

- Average bit-rate. Denoting the mean video quality over the entire session.
- Bit-rate stability. Describing the jittering of video quality during the entire session.
- Start-up delay. Denoting the waiting duration between the time that the user requests streaming service and the time that media player starts to play.

Paper [7] pointed out that the buffering ratio is most critical across genres. For example, 1% increase in buffering reduces 3 minutes for a 90-minutes live video streaming. [16] showed that the total time spent rebuffering and the frequency of rebuffering events have substantial impact on QoE. Under this context, media servers and network operators face a crucial challenge on how to avoid the degradation of user perceived media quality based on these metrics. However, development of new models as function of these metrics can help operators and content publishers to better invest their network and server resources toward optimizing these metrics that really matter for QoE.

In this paper we focus on a setting in which a video is streamed over a wireless network which is subject to a lot of constraints like bandwidth limitation and rate fluctuations due to the frequent changes of channel states and mobility [5]. Indeed, time varying network capacity is especially relevant when considering wireless networks where such variations can be caused by fast fading and slow fading due to shadowing, dynamic interference, and changing loads. To address this issue, we focus on performance modelling and analysis of a streaming video with Long Range Dependence (LRD) traffic and variable service capacity. Due to the inherent difficulty and complexity of modelling fractal-like LRD traffic, we assume that the arrival of packets at player buffer are characterized by a Markov Modulated Fluid Model, which accurately approximates the traffic exhibiting LRD behaviour and mimics the real behaviour of multimedia traffic with short-term and long-term correlation [10]. In comparison to related works, our whole analysis is on transient regime. We construct sets of Partial Differential Equations (PDEs) to derive the starvation probability generating function using the external environment,

which is described by the Continuous Time Markov Chain (CTMC). This approach predicts the starvation probability as function of the file size as well as the prefetching threshold. Moreover we provide relevant results to understand on how the starvation probabilities are impacted by the variation of traffic load and prefetching threshold. We do simulations to show the accuracy of our model using ns-3. Achieving this goal, we are able to identify through our model the dependencies between quality metrics. For example, start-up delay can reduce rebuffering ratio. Similarly bitrate rate switching can reduce buffering.

With the results developed in this work, we are able to answer the fundamental questions: How many frames should the media player prefetch to optimize the users' quality of experience? From what file size the adaptive coding is relevant to avoid the starvation? How bit-rate switching impacts the QoE metrics? Knowing these answers enables the user to maximize his QoE realising the tradeoffs among different metrics incorporating user preferences on rebuffering ratio, start-up delay and quality [13]–[15]. We further introduce an optimization problem which takes these key factors in order to achieve the optimal tradeoff between them. We adopt a more flexible method by defining an objective of QoE by associated a weight for each metric based on user preferences [13].

II. RELATED WORKS

QoE analysis over wireless networks has been studied for many years. In [1], authors study the QoE in a shared fast-fading channel using an analytical framework based on Takacs Ballot theorem. They use a GI/D/1 queue to model the system, so they assume that the arrival process is independent and identically distributed (i.i.d). In [9], the analysis of buffer starvation using M/M/1 queue is performed. They use a recursive method to compare the results with the Ballot theorem method even if the recursive method did not offer explicit results. They assume an i.i.d arrival process that is a rough model of streaming services over the wireless networks. Since the performance measures depend on the autocorrelation structure of the traffic, a consensus exists about the limitation of the Poisson process to model the traffic behaviour. In [17], authors develop an analytical framework to investigate the impact of network dynamics on the user perceived video quality, they model the playback buffer by a G/G/1 queue and use the diffusion approximation method to compute the QoE. The QoE of streaming from the perspective of the network flow dynamics is studied in [5]. The throughput of a tagged user is governed by the number of the other users in the network. This study shows that the network flow dynamics is the fundamental reason for playback starvation.

The rest of this paper is organized as follows: In section III, we describe the system model while section IV presents the analysis of the queuing system model. Section V describes the performance analysis of the quality of experience and section VI presents explicit results for two states MMFM. Section VII shows numerical results and section VIII concludes this paper.

III. SYSTEM MODEL DESCRIPTION

We consider a single user receiving a media file with finite size Z in streaming. Generally, media files are divided into blocks of frames. When a user makes a request the server segments this media into frames and transfers them to the user through the network (wired and wireless links). When frames traverse the internet, their arrivals are not deterministic due to the dynamics of the available bandwidth. One of the main characteristics of wireless traffic and Internet traffic in general is the rate fluctuation caused by fast fading and slow fading due to shadowing, dynamic interference, and changing load. Moreover, data packet arrivals in cellular networks are found to be correlated over both short and long-time scales. This is generally due to the arrival of packets bursts of comparable size, often leading to high instantaneous arrival rates. Hence, video flows through the Internet with fluctuating speed. In this paper, we assume that frames arrive to the play-out buffer with a rate that can take values from finite set $S = \{\lambda_i, i = 1, 2, \dots, L\}$. The rate of arrival frames is governed by a Continuous-Time Markov Chain (CTMC) $\{I(t), t \geq 0\}$ with infinitesimal generator Q .

$$Q = \begin{pmatrix} q_{1,1} & q_{1,2} & \cdots & q_{1,L} \\ q_{2,1} & q_{2,2} & \cdots & q_{2,L} \\ \vdots & \vdots & \ddots & \vdots \\ q_{L,1} & q_{L,2} & \cdots & q_{L,L} \end{pmatrix}$$

where $q_{i,i} = -\sum_{j \neq i} q_{ij}$.

The maximum buffer size is assumed to be large enough so that the whole file can be stored. At the user side, incoming frames are stocked in a buffer and from there they are played with a rate μ (e.g., 25 frames per second (fps)) in the TV and movie-making business. We quantify the user perceived media quality using two measures called start-up delay and starvation. There is ongoing research on mapping these two measures on standard human evaluated QoE measures. As explained earlier, the media player wants to avoid the starvation by prefetching packets. However, this action might incur a long waiting time. In what follows, we reveal the relationship between the start-up delay and the starvation behaviour, with the consideration of the file size.

We consider a fluid model that has been proven to be a powerful modeling paradigm in many applications and relevant to capture the key characteristics that determine the performance of networks. Let $r_i = \lambda_i - \mu$ denote the effective input rate in state i . Hence the matrix of the effective rates is R , which is a diagonal matrix $diag\{\lambda_1 - \mu, \lambda_2 - \mu, \dots, \lambda_L - \mu\}$. We denote by $X(t)$ the length of playout buffer of playback at time t . Let τ be the first time the buffer is empty before reaching the end of the file, i.e., $\tau = \inf\{t > 0 : X(t) = 0\}$ and T_x be the start-up delay where x is the prefetching threshold. In the next section we provide mathematical analysis to compute the distribution of the number of starvation and start-up delay for a general bursty arrival process.

IV. ANALYSIS OF THE QUEUING SYSTEM MODEL

A. Laplace Transform of the Starvation Probability

We compute the Laplace transform of the probability of starvation given the Continuous-Time Markov Chain $\{I(t), t \geq 0\}$. We define $H_{ij}(x, t)$ to be the probability of starvation in state j before time t , given the initial state i and the initial queue length x .

$$H_{ij}(x, t) = P\{\tau \leq t, I(\tau) = j | X(0) = x, I(0) = i\} \quad (1)$$

for $i, j = 1, 2, \dots, L$, $x > 0$ and $t \geq 0$. It is clear that the CTMC cannot be in a state j at time τ if $r_j > 0$. Hence

$$H_{ij}(x, t) = 0 \quad \text{for all } t \geq 0, \quad x \geq 0 \quad \text{if } r_j > 0.$$

$$H_{ij}(x, t) = 0 \quad \text{for all } t \geq 0, \quad x > 0 \quad \text{if } r_j = 0.$$

$$H_{ij}(0, t) = 1 \quad \text{for all } t \geq 0, \quad \text{if } r_j = 0.$$

Let $\pi = (\pi_1, \pi_2, \dots, \pi_L)$ be the steady state probability vector of the CTMC $\{I(t), t \geq 0\}$ where π_i is the probability to be in the state i at the stationary regime. The expected input and output rates are $\sum_{i \in S} \pi_i \lambda_i$ and $\sum_{i \in S} \pi_i \mu_i$ respectively. The buffer is stable if $\sum_{i \in S} \pi_i \lambda_i < \sum_{i \in S} \pi_i \mu_i$. Conditioning on the first transition from the state i at time 0 we have

$$H_{ij}(x, t) = \sum_{k \neq i} q_{ik} \Delta t H_{kj}(x + r_i \Delta t, t - \Delta t) + (q_{ii} \Delta t + 1) H_{ij}(x + r_i \Delta t, t - \Delta t) + o(\Delta t) \quad (2)$$

Taking the limit $\lim_{\Delta t \rightarrow 0} \frac{H_{ij}(x, t) - H_{ij}(x, t - \Delta t)}{\Delta t}$ and after some algebraic simplification we obtain the following partial differential equation

$$\frac{\partial H_j(x, t)}{\partial t} - R \frac{\partial H_j(x, t)}{\partial x} = Q H_j(x, t) \quad (3)$$

with the initial conditions

$$H_{ij}(0, t) = \begin{cases} 1 & \text{if } i = j \text{ and } t \geq 0 \\ 0 & \text{if } i \neq j \text{ and } r_i < 0 \end{cases}$$

$$H_{ij}(x, 0) = 0 \quad \text{for all } i \neq j \text{ and } x \geq 0,$$

$$H_{jj}(x, 0) = 0 \quad \text{for } x > 0.$$

where $H_j(x, t) = [H_{1j}(x, t), H_{2j}(x, t), \dots, H_{Lj}(x, t)]$. The Laplace Stieljes Transform (LST) of $H_{ij}(x, t)$ is

$$\begin{aligned} \tilde{H}_{ij}(x, \omega) &= \int_0^\infty e^{-\omega t} dH_{ij}(x, t) \\ &= E[e^{-\omega \tau}; I(\tau) = j | X(0) = x, I(0) = i] \end{aligned}$$

for $i, j = 1, \dots, L$ and $\tilde{H}_j(x, \omega) = [\tilde{H}_{1j}(x, \omega), \dots, \tilde{H}_{Lj}(x, \omega)]$. Taking the LST of Equation (3) and using the fact that $H_{ij}(x, 0) = 0$ for all $x > 0$, we find

$$R \frac{d\tilde{H}_j(x, \omega)}{dx} = (\omega I - Q) \tilde{H}_j(x, \omega) \quad (4)$$

For a fixed value of ω , we take

$$\tilde{H}_j(x, \omega) = e^{s_k(\omega)x} \phi(\omega)$$

as a solution to Equation (4). Substituting in (4) we get

$$R s(\omega) \phi(\omega) = (\omega I - Q) \phi(\omega)$$

where the scalar $s(\omega)$ and the vector $\phi(\omega)$ are to be determined. The theorem 3.3 from [6] gives

$$\tilde{H}_j(x, \omega) = \sum_{s_k(\omega) \in I^-} a_{kj} e^{s_k(\omega)x} \phi_j^k(\omega) \quad (5)$$

where the coefficients a_{kj} are obtained by solving

$$\sum_{s_k(\omega) \in I^-} a_{kj} \phi_j^k(\omega) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j, r_i \leq 0 \end{cases}$$

$s_k(\omega)$ are the roots with negative real parts of $\Delta(s, \omega) = \det(Q + sR - \omega I)$ and $\phi^k(\omega)$ are the corresponding eigenvectors satisfying the equation

$$(Q + s(\omega)R - \omega I) \phi(\omega) = 0 \quad (6)$$

B. Laplace Transform of the Start-up delay

We consider the previous system during the prefetching process and we denote by $X_s(t)$ the length playout buffer of playback at time t . Let

$$T_x = \inf\{t \geq 0 : X_s(t) \geq x\}$$

be the first time that the length playout buffer reaches x . T_x is the time that the system will take to accumulate x content in the buffer. This distribution is difficult to solve directly, so we resort to the following duality problem:

Duality problem: What is the starvation probability by time t if the queue is depleted with rate λ_i and the duration of prefetching contents is x ?

This duality problem allows us to compute the prefetching delay as a probability of starvation. We define $U_{ij}(x, t)$ to be the probability of starvation before time t at the state j , conditioning on the initial state i and the initial prefetching content x , i.e., the start-up threshold.

$$U_{ij}(x, t) = P\{T_x \leq t, I(T_x) = j | I(0) = i, X_s(0) = x\} \quad (7)$$

for $i, j = 1, 2, \dots, L$, $x > 0$ and $t \geq 0$.

Conditioning on the first transition from the state at time 0,

$$U_{ij}(x, t) = \sum_{k \neq i} q_{ik} \Delta t U_{kj}(x - \lambda_i \Delta t, t - \Delta t) + (q_{ii} \Delta t + 1) U_{ij}(x - \lambda_i \Delta t, t - \Delta t) + o(\Delta t) \quad (8)$$

Taking the limit $\lim_{\Delta t \rightarrow 0} \frac{U_{ij}(x, t) - U_{ij}(x, t - \Delta t)}{\Delta t}$ and after some algebraic simplification we obtain the following partial differential equation

$$\frac{\partial U_j(x, t)}{\partial t} - R \frac{\partial U_j(x, t)}{\partial x} = Q U_j(x, t) \quad (9)$$

with the same initial conditions as in section IV-A, where $R = \text{diag}\{-\lambda_1, -\lambda_2, \dots, -\lambda_L\}$ and $U_j(x, t) = [U_{1j}(x, t), U_{2j}(x, t), \dots, U_{Lj}(x, t)]$.

$$\tilde{U}_j(x, \omega) = \sum_{s_k(\omega)} a_{kj} e^{s_k(\omega)x} \phi_j^k(\omega) \quad (10)$$

where the coefficients a_{kj} are obtained by solving

$$\sum_{s_k(\omega)} a_{kj} \phi_j^k(\omega) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

and $s_k(\omega)$ are the roots of $\det(Q + sR - \omega I)$ and $\phi^k(\omega)$ are the corresponding eigenvectors.

In what follows, we compute the probability that the prefetching ends at a given state [3]. For this purpose, we define

$$V_{ij}(q, x) = P\{I(T_x) = j | I(0) = i, X_s(0) = q\} \quad (11)$$

to be the probability that the prefetching ends at state j given the initial state i and the initial queue length q where x is the prefetching threshold. In the time interval $[0, h]$, conditioning on the first transition from the state at time 0, we have

$$V_{ij}(q, x) = (1 + q_{ii}h)V_{ij}(q + \lambda_i h, x) + \sum_{k \neq i} q_{ik}hV_{kj}(q + \lambda_i h, x) + o(h) \quad (12)$$

After some algebraic simplification and letting $h \rightarrow 0$ yields the differential equation

$$\text{diag}\left\{\frac{1}{\lambda_i}\right\} \dot{\mathbf{V}}(q, x) = -Q\mathbf{V}(q, x) \quad (13)$$

with the boundary condition

$$V_{ij}(x, x) = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Let

$$Q_v = \text{diag}\left\{\frac{1}{\lambda_i}\right\} \cdot (-Q)$$

Eq. (13) becomes: $\dot{\mathbf{V}}(q, x) = Q_v \mathbf{V}(q, x)$. $\mathbf{V}(q, x)$ is given by

$$\mathbf{V}(q, x) = \exp(Q_v q) \cdot \mathbf{V}(0, x) \quad (14)$$

Using Eq.(14) and the initial conditions, we get

$$\mathbf{V}(q, x) = D_v \exp(\Lambda_v(q - x)) D_v^{-1} \cdot \mathbf{V}(x, x) \quad (15)$$

where $D_v \cdot \Lambda_v \cdot D_v^{-1} = Q_v$, Λ_v is the diagonal matrix containing all the eigenvalues of Q_v and D_v is an invertible matrix.

C. The Probability of Starvation and the Start-up Delay

In the previous sections we derived explicit expressions for the Laplace-Stieltjes Transform of the probability of starvation and the start-up delay. In this section, we present theoretical models to find the corresponding probability of starvation and start-up delay. The Laplace Stieljes Transform of $H_{ij}(x, t)$ is

$$\begin{aligned} \tilde{H}_{ij}(x, \omega) &= E[e^{-\omega\tau}; I(\tau) = j | X(0) = x, I(0) = i] \\ &= \int_0^\infty e^{-\omega t} dH_{ij}(x, t) = \int_0^\infty e^{-\omega t} h_{ij}(x, t) \end{aligned}$$

where h_{ij} is the probability density function of H_{ij} .

Lemma 1 (Bromwich inversion integral). *Given the Laplace transform \tilde{h} , the function value $h(t)$ can be recovered from the contour integral*

$$h(t) = \frac{1}{2\pi i} \int_{b-i\infty}^{b+i\infty} e^{\omega t} \tilde{h}(\omega) d\omega, \quad t > 0, \quad (16)$$

where b is a real number to the right of all singularities of \tilde{h} , $i^2 = -1$, and the contour integral yields the value 0 for $t < 0$.

It is shown in [8] that for real value functions, h has the following form

$$h(t) = \frac{2e^{bt}}{\pi} \int_0^\infty \text{Re}(\tilde{h}(b + iu)) \cos(ut) du \quad (17)$$

According to the *Bromwich inversion integral*, $h(t)$ can be calculated from the transform \tilde{h} by performing a numerical integration (quadrature). We use a specific algorithm based on the Bromwich inversion integral. It is based on a variant of the Fourier-series method - the trapezoidal rule - which proves to be remarkably effective. If we use a step size T , then the trapezoidal rules gives

$$\begin{aligned} h(t) \approx h_T(t) &\equiv \frac{T e^{bt}}{\pi} \text{Re}(\tilde{h}(b)) \\ &+ \frac{2T e^{bt}}{\pi} \sum_{k=1}^\infty \text{Re}(\tilde{h}(b + ikh)) \cos(kht) \end{aligned} \quad (18)$$

where $\text{Re}(\tilde{h}(b)) = \tilde{h}(b)$ since b is real. Replacing $\tilde{h}(\omega)$ by $\frac{\tilde{H}_{ij}(x, \omega)}{\omega}$ which is the Laplace transform of $H_{ij}(x, t)$, we get the probability of starvation before time t

$$\begin{aligned} P\{\tau \leq t\} &= \frac{2T e^{bt}}{\pi} \left[\frac{\tilde{H}_{ij}(x, b)}{2b} \right. \\ &\left. + \sum_{k=1}^\infty \text{Re}\left(\frac{\tilde{H}_{ij}(x, b + ikh)}{b + ikh}\right) \cos(kht) \right] \end{aligned} \quad (19)$$

The infinite series in (19) can simply be calculated by simple truncating because it converges, but more efficient algorithm can be obtained by applying a summation acceleration method. An acceleration technique that has proven to be effective in our context is Euler summation, after transforming the infinite sum into a nearly alternating series in which successive summands alternate in sign. We convert (19) into a nearly alternating series by letting $T = \pi/lt$ and $b = A/2lt$

$$\begin{aligned} h_T(t) \equiv h_{A,l}(t) &= \frac{e^{A/2lt}}{2lt} + \frac{2e^{A/2lt}}{lt} \sum_{k=1}^\infty \tilde{h}\left(\frac{A}{2lt} + \frac{ik\pi}{lt}\right) e^{ik\pi/l} \\ &= \sum_{k=0}^\infty (-1)^k a_k(t) \end{aligned}$$

where

$$a_k(t) = \frac{e^{A/2l}}{2lt} \left(\tilde{h}\left(\frac{A}{2lt}\right) 1_{\{k=0\}} + 2 \sum_{j=1}^l \operatorname{Re} \left[\tilde{h}\left(\frac{A}{2lt} + \frac{ij\pi}{lt} + \frac{ik\pi}{t}\right) e^{ij\pi/l} \right] \right)$$

Let s_n be the approximation $h_{A,l}(t)$ with the infinite series truncated to n terms, i.e.,

$$s_n = \sum_{k=0}^n (-1)^k a_k$$

where t is suppressed in the notation and $a_k \equiv a_k(t)$. We apply Euler summation to m terms after an initial n , so that the Euler sum approximation is

$$E(m, n) \equiv E(m, n, t) \equiv \sum_{k=0}^m C_m^k 2^{-m} s_{n+k} \quad (20)$$

Euler summation can be very simply described as the weighted average of the last m partial sums by a binomial probability with parameter m and $p = 1/2$. Hence, (20) is the binomial average of the terms $s_n, s_{n+1}, \dots, s_{n+m}$. The implementation of the algorithm takes into account the values of (l, m, n, A) . As in [8], we use $(1, M, M, 2\ln(10)M/3)$ where $M = 64$. After simplification and letting l be 1, we get the Euler approximation $E(t)$ of the inverse $h(t)$ which seems to be a good approximation

$$E(t) = \sum_{k=0}^m C_m^k 2^{-m} \sum_{q=0}^{n+k} \frac{e^{A/2}}{2t} \left[\tilde{h}\left(\frac{A}{2t}\right) 1_{\{q=0\}} - 2 \operatorname{Re} \left(\tilde{h}\left(\frac{A}{2t} + \frac{i\pi(1+q)}{t}\right) \right) \right] \quad (21)$$

Eq. (21) looks complicated, but it consists of only $((m+1)(m+2n+2))/2$ additions, that is a low computation level. To have the cdf H_{ij} , we just replace \tilde{h} by $\frac{\tilde{H}_{ij}(t)}{t}$. The same formula holds for the start-up delay distribution in replacing \tilde{h} by $\frac{\tilde{U}_{ij}(t)}{t}$.

V. PERFORMANCES ANALYSIS OF THE QUALITY OF EXPERIENCE

In this section we compute the QoE metrics based on the analysis derived in the previous section.

A. The Probability of Starvation

We consider a single user receiving a media file with size Z . The necessary time to play the whole video if there is no starvation is $\frac{Z}{\mu}$. Hence, using the first passage time distribution $H_{ij}(x, t)$, the probability of starvation happened in state j before reaching the end of file given the initial state i , is given

by

$$P_s = \sum_{k=0}^m C_m^k 2^{-m} \sum_{q=0}^{n+k} e^{A/2} \left[\frac{\tilde{H}_{ij}\left(\frac{\mu A}{2Z}\right)}{A} 1_{\{q=0\}} - 2 \operatorname{Re} \left(\frac{\tilde{H}_{ij}\left(\frac{\mu A + 2i\pi\mu(1+q)}{2Z}\right)}{A + 2i\pi(1+q)} \right) \right] \quad (22)$$

The starvation of probability before time t gives an idea of the severity of the starvation during the video session. Let $D_{ij}(x) := E[\tau, I(\tau) = j | \tau < \infty, I(0) = i, X(0) = x]$ be the mean continuous playback time if the initial state is i , the prefetching threshold is x and the starvation happens in state j . $D_{ij}(x)$ is an important measure for the severity of starvations. A small $D_{ij}(x)$ means that the starvation events happen frequently. We find $D_{ij}(x)$ by taking derivatives of $\tilde{H}_{ij}(x, \omega)$ in $\omega = 0$.

$$D_{ij}(x) = - \frac{\partial \tilde{H}_{ij}(x, \omega)}{\partial \omega} \Big|_{\omega=0} \quad i, j = 1, \dots, L$$

When the user starts the video session, the initial state is unknown to the system. The video starts playing when the prefetching process is finished. Conditioning on the distribution of the entry states π , the distribution of the states that the playback process begins (or prefetching process ends) is computed by $\pi \cdot \mathbf{V}(0, x)$. Recalling that $V_{ij}(0, x)$ is the probability that the prefetching phase ends at state j knowing that the video session starts at state i . Then the starvation probability with the prefetching threshold x is obtained by

$$P_s(x) = \pi \cdot \mathbf{V}(0, x) \cdot \mathbf{H}\left(x, \frac{Z}{\mu}\right) \quad (23)$$

where \mathbf{H} is a column vector, $\mathbf{H} = (H_1, H_2, \dots, H_L)^T$ and $H_i = \sum_{j=1}^L H_{ij}$. $P_s(x)$ is called the overall starvation probability. The probability of no starvation is $1 - P_s(x)$.

B. The distribution of the Start-up delay

The start-up delay is proportional to the start-up threshold. But in the QoE literature, it is more practical to consider the delay rather than the threshold because the delay has a direct impact on the streaming user behaviour. Using the results of the sections IV-C and IV-B, we derive the cumulative distribution function of the start-up delay

$$U_{ij}(x, t) = \sum_{k=0}^m C_m^k 2^{-m} \sum_{q=0}^{n+k} e^{A/2} \left[\frac{\tilde{U}_{ij}\left(\frac{\mu A}{2Z}\right)}{A} 1_{\{q=0\}} - 2 \operatorname{Re} \left(\frac{\tilde{U}_{ij}\left(\frac{\mu A + 2i\pi\mu(1+q)}{2Z}\right)}{A + 2i\pi(1+q)} \right) \right] \quad (24)$$

where x is the start-up threshold, Z is the file size and m, n, A are the Euler Summation Algorithm parameters.

C. The generating function of the starvation events

When a starvation event happens, the media player pauses until x contents are re-buffered. We are interested in the probability distribution of the starvations, given the file size Z . We define a *path* as a complete sequence of frames arrivals

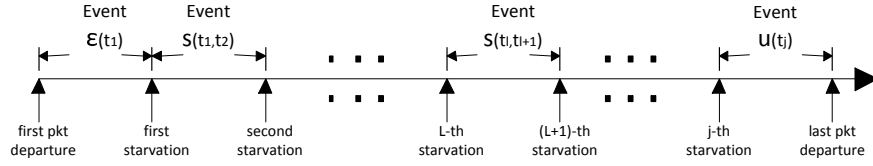


Fig. 1. A path with j starvations

and departures. We illustrate a typical path with j starvations in Fig. 1. The path can be decomposed into three types of mutually exclusive events as follows:

- Event $\mathcal{E}(t_1)$: the buffer becoming empty for the first time in the entire path.
- Event $\mathcal{S}_l(t_l, t_{l+1})$: the empty buffer after the instant t_{l+1} given that the previous empty buffer happens at t_l .
- Event $\mathcal{U}_j(t_j)$: the last empty buffer observed after the instant t_j .

Obviously, a path with j starvations is composed of a succession of events

$$\mathcal{E}(t_1), \mathcal{S}_1(t_1, t_2), \mathcal{S}_2(t_2, t_3), \dots, \mathcal{S}_{j-1}(t_{j-1}, t_j), \mathcal{U}_j(t_j)$$

We let $P_{\mathcal{E}(t_1)}$, $P_{\mathcal{S}_l(t_l, t_{l+1})}$ and $P_{\mathcal{U}_j(t_j)}$ be the probabilities of events $\mathcal{E}(t_1)$, $\mathcal{S}_l(t_l, t_{l+1})$ and $\mathcal{U}_j(t_j)$ respectively. The probability distribution of event $\mathcal{E}(t_1)$ is expressed as

$$P_{\mathcal{E}(t_1)} = \begin{cases} 0, & \text{if } \mu t_1 < x \text{ or } \mu t_1 \geq Z; \\ \pi \cdot \mathbf{V}(0, x) \cdot \mathbf{h}(x, t_1), & \text{otherwise.} \end{cases} \quad (25)$$

where \mathbf{V} and \mathbf{h} are $L \times L$ and $L \times 1$ matrices respectively. The first starvation cannot happen at the departure of first $(x-1)$ contents because of the prefetching of x contents. It cannot happen after all Z contents have been served because this empty buffer is not a starvation. For $\mu t_1 \in [x, Z[$ the starvation happens at time t_1 conditioned on the states that the playback process begins. The probability distribution of event $\mathcal{U}_j(t_j)$ is given by

$$P_{\mathcal{U}_j(t_j)} = \begin{cases} 0, & \text{if } \mu t_j < jx \text{ or } \mu t_j \geq Z; \\ 1, & \text{if } Z - x \leq \mu t_j < Z; \\ \mathbf{V}(0, x) \cdot (\mathbf{1} - \mathbf{H}(x, \frac{Z}{\mu} - t_j)), & \text{otherwise.} \end{cases} \quad (26)$$

where \mathbf{H} is a column vector. t_j is the time of the j -th starvation. The extreme case is that these j starvations take place consecutively. Then μt_j should be greater than jx . Otherwise there cannot have j starvations. If μt_j is no less than $Z - x$, the media player resumes until all the remaining content $Z - \mu t_j$ is stored in the buffer. Then, starvation will not appear afterwards. In the remaining case, it is the probability of having no starvation after time t_j . We denote by $P_s(j)$ the probability of having j starvations. The case with one starvation is given by

$$P_s(1) = \int_{t=0}^{\frac{Z}{\mu}} P_{\mathcal{E}(t)} \cdot P_{\mathcal{U}_1(t)} dt \quad (27)$$

To compute the probability of having more than one starvation, we need to find the probability of event $\mathcal{S}_l(t_l, t_{l+1})$. μt_l should

not be less than lx in order to have l starvations. Given that the buffer is empty just after time t_l , the $(l+1)^{th}$ starvation cannot happen at $\mu t_{l+1} \in [\mu t_l + 1, \mu t_l + x - 1]$ because of the prefetching process. Since there are j starvations in total, the $(l+1)^{th}$ starvation must satisfy $\mu t_{l+1} < Z - (j-l-1)x$. We next compute the remaining case that the l^{th} and the $(l+1)^{th}$ starvations happen at time t_l and t_{l+1} respectively. We compute this probability using the first passage time density when the starvation happens at time t_{l+1} and the initial time was t_l with a prefetching process. $P_{\mathcal{S}_l(t_l, t_{l+1})}$ is expressed as

$$\begin{cases} \mathbf{V}(0, x) \cdot \mathbf{h}(x, t_{l+1} - t_l), \\ \text{if } \mu t_l \geq lx, \mu t_l + x \leq \mu t_{l+1} < Z - (j-l-1)x; \\ 0, & \text{otherwise.} \end{cases} \quad (28)$$

We use in this method a trick that concerns the time scale. Every time the player resumes for the prefetching process we resume also the time scale, that means if the starvation happens at time t , the player will start playing at the same time t with x initial contents in the buffer. The probability of having j ($j \geq 2$) starvations is given by

$$P_s(j) = \int_{t_1=0}^{\frac{Z}{\mu}} \int_{t_2=0}^{\frac{Z}{\mu}} \dots \int_{t_{j-1}=0}^{\frac{Z}{\mu}} \int_{t_j=0}^{\frac{Z}{\mu}} P_{\mathcal{E}(t_1)} \cdot P_{\mathcal{S}_1(t_1, t_2)} \dots P_{\mathcal{S}_{j-1}(t_{j-1}, t_j)} \cdot P_{\mathcal{U}_j(t_j)} dt_1 dt_2 \dots dt_{j-1} dt_j \quad (29)$$

In the next section, we provide explicit expressions of QoE metrics where CTMC has two states.

VI. THE 2-STATE MMFM SOURCE

In this section, we consider a special case in which the CTMC has two states : $\{1, 2\}$ (see Fig. 2) with infinitesimal generator Q and rate matrix R

$$Q = \begin{pmatrix} -\beta & \beta \\ \alpha & -\alpha \end{pmatrix} \quad R = \begin{pmatrix} \lambda_2 - \mu & 0 \\ 0 & \lambda_1 - \mu \end{pmatrix}$$

Our objective is to understand the interaction between the

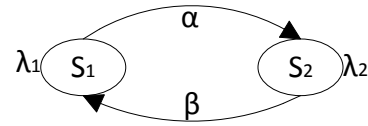


Fig. 2. The two-state MMPP source

parameters of arrival process and the probability of starvation.

Using the results of section IV-A irrespective of the condition of stability of the queue $\pi_1\lambda_1 + \pi_2\lambda_2 < \mu$, we get:

$$\begin{aligned}\Delta(s, \omega) &= \det(Q + sR - \omega I) \\ &= (\lambda_1 - \mu)(\lambda_2 - \mu)s^2 - [(\lambda_1 - \mu)(\omega + \beta) \\ &\quad + (\lambda_2 - \mu)(\omega + \alpha)]s + \omega(\omega + \alpha + \beta)\end{aligned}\quad (30)$$

It is a polynomial of degree 2 in s where the two zeros are given by:

$$s_1(\omega) = \frac{b + \sqrt{b^2 - 4\omega(\omega + \alpha + \beta)(\lambda_1 - \mu)(\lambda_2 - \mu)}}{2(\lambda_1 - \mu)(\lambda_2 - \mu)} \quad (31)$$

$$s_2(\omega) = \frac{b - \sqrt{b^2 - 4\omega(\omega + \alpha + \beta)(\lambda_1 - \mu)(\lambda_2 - \mu)}}{2(\lambda_1 - \mu)(\lambda_2 - \mu)} \quad (32)$$

where $b = (\lambda_1 - \mu)(\omega + \beta) + (\lambda_2 - \mu)(\omega + \alpha)$. Equation (5) contains terms with only $\text{Re}(s_k(\omega)) < 0$. So we have to determine the signs of $\text{Re}(s_1(\omega))$ and $\text{Re}(s_2(\omega))$. The next propositions give the placement of these two zeros in the complex plane.

Proposition 1. Let $\lambda_1 > \lambda_2$,

1. $\lambda_2 > \mu \Rightarrow \lambda_1 > \mu$, so $\lambda_1 - \mu > 0$ and $\lambda_2 - \mu > 0$ then no starvation.
2. $\lambda_2 < \mu$ and $\lambda_1 > \mu$, so $\lambda_2 - \mu < 0$ and $\lambda_1 - \mu > 0$ then $\text{Re}(s_2(\omega)) > 0$ and $\text{Re}(s_1(\omega)) < 0$.
3. $\lambda_2 < \mu$ and $\lambda_1 < \mu$, so $\lambda_2 - \mu < 0$ and $\lambda_1 - \mu < 0$ then $\text{Re}(s_2(\omega)) < 0$ and $\text{Re}(s_1(\omega)) < 0$.
4. $\lambda_2 = \mu \Rightarrow \lambda_1 > \mu$ then no starvation.
5. $\lambda_1 = \mu \Rightarrow \lambda_2 < \mu$, $s_2(\omega) = s_1(\omega) = s(\omega) = \frac{\omega(\omega + \alpha + \beta)}{(\lambda_2 - \mu)(\omega + \alpha)}$ and $\text{Re}(s(\omega)) < 0$.

Proposition 2. Let $\lambda_2 > \lambda_1$,

1. $\lambda_1 > \mu \Rightarrow \lambda_2 > \mu$, so $\lambda_1 - \mu > 0$ and $\lambda_2 - \mu > 0$ then no starvation.
2. $\lambda_1 < \mu$ and $\lambda_2 > \mu$, so $\lambda_1 - \mu < 0$ and $\lambda_2 - \mu > 0$ then $\text{Re}(s_2(\omega)) > 0$ and $\text{Re}(s_1(\omega)) < 0$.
3. $\lambda_1 < \mu$ and $\lambda_2 < \mu$, so $\lambda_1 - \mu < 0$ and $\lambda_2 - \mu < 0$ then $\text{Re}(s_2(\omega)) < 0$ and $\text{Re}(s_1(\omega)) < 0$.
4. $\lambda_1 = \mu \Rightarrow \lambda_2 > \mu$ then no starvation.
5. $\lambda_2 = \mu \Rightarrow \lambda_1 < \mu$, $s_2(\omega) = s_1(\omega) = s(\omega) = \frac{\omega(\omega + \alpha + \beta)}{(\lambda_1 - \mu)(\omega + \beta)}$ and $\text{Re}(s(\omega)) < 0$.

Proposition 3. Let $\lambda_1 = \lambda_2 = \lambda$,

1. $\lambda > \mu$, no starvation.
2. $\lambda < \mu$, $\text{Re}(s_2(\omega)) < 0$ and $\text{Re}(s_1(\omega)) < 0$.
3. $\lambda = \mu$, no starvation because of the prefetching.

The LST $\tilde{H}_j(x, \omega)$ of the distribution is given in the next theorem.

Theorem 1. 1. When $\lambda_1 < \mu$, $\lambda_2 \geq \mu$, $\tilde{H}_2(x, \omega) = 0$ and

$$\tilde{H}_1(x, \omega) = \begin{bmatrix} \tilde{H}_{11}(x, \omega) \\ \tilde{H}_{21}(x, \omega) \end{bmatrix} = e^{s_1(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{\beta} \end{bmatrix}$$

2. When $\lambda_2 < \mu$, $\lambda_1 \geq \mu$, $\tilde{H}_1(x, \omega) = 0$ and

$$\tilde{H}_2(x, \omega) = \begin{bmatrix} \tilde{H}_{22}(x, \omega) \\ \tilde{H}_{12}(x, \omega) \end{bmatrix} = e^{s_2(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_2(\omega)}{\beta} \end{bmatrix}$$

3. When $\lambda_1 < \mu$, $\lambda_2 < \mu$,

$$\tilde{H}_1(x, \omega) = a_{21}e^{s_2(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_2(\omega)}{\beta} \end{bmatrix} + a_{11}e^{s_1(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{\beta} \end{bmatrix}$$

$$\tilde{H}_2(x, \omega) = a_{22}e^{s_2(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_2(\omega)}{\beta} \end{bmatrix} + a_{12}e^{s_1(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{\beta} \end{bmatrix}$$

where

$$\begin{aligned}a_{11} &= \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{(\lambda_2 - \mu)(s_2(\omega) - s_1(\omega))}, a_{21} = \frac{\beta + \omega - (\lambda_2 - \mu)s_2(\omega)}{(\lambda_2 - \mu)(s_1(\omega) - s_2(\omega))} \\ a_{22} &= \frac{\beta}{(\lambda_2 - \mu)(s_1(\omega) - s_2(\omega))}, a_{12} = \frac{\beta}{(\lambda_2 - \mu)(s_2(\omega) - s_1(\omega))}\end{aligned}$$

The proof of this theorem can be find in the appendix of [18]. Taking $\lambda_2 = 0$ gives the first passage time distribution for the ON-OFF source.

VII. NUMERICAL ANALYSIS

A. Simulation

We use ns-3 simulator in order to compare the dynamics of the process with our model. The simulation topology consists on a server and a client in order to simulate the queue model. The server sends the traffic to the client following the Continuous Time Markov Chain. The client holds a buffer where the traffic is stored. The parameters of the traffic depend on the CTMC parameters. Then we analyze the behavior of the client buffer content which simulates the player. We run 10 simulations and compute the 95% confidence interval on all observed metrics, but it is not shown on all the figures for improving readability because it is very narrow. We first show the accuracy of the method that we use to invert the Laplace Transform. In Fig. 3, we plot the known inverse Laplace Transform of the function $f(t) = e^{-2t}\sin(\pi t)$ that is $f(s) = \pi/((s + \pi)^2 + \pi^2)$ and the inverse using formula (21) of section IV-C. Fig. 4 shows the starvation probability for a two states MMFM source for $\lambda_1 = 2$, $\lambda_2 = 30$ and $\mu = 25$, that means the buffer size increases on state 2 and decreases on state 1. This is done for states transitions $\alpha = 2$ and $\beta = 6$. Fig. 6 illustrates the impact of the start-up threshold x on the probability of no starvation. Fig. 7 shows the probability of having one starvation. These simulation results validate the correctness of our analysis. Hence, in the following experiments, we only illustrate the analytical results.

B. Performances Evaluation

Fig. 8 gives the CDF of the start-up delay for different values of start-up threshold x . We can see that the start-up delay increases with x . On the other hand figure 6 illustrates the impact of the start-up threshold x on the probability of no starvation. When x is large enough (near 300 pkts in the

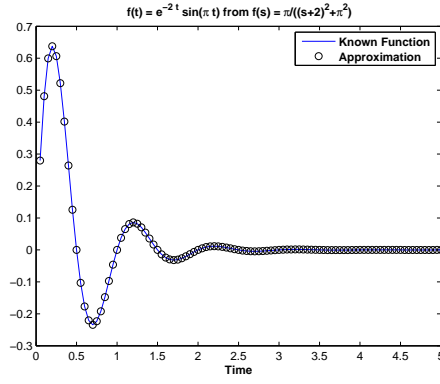


Fig. 3. The accuracy of Euler Summation Algorithm

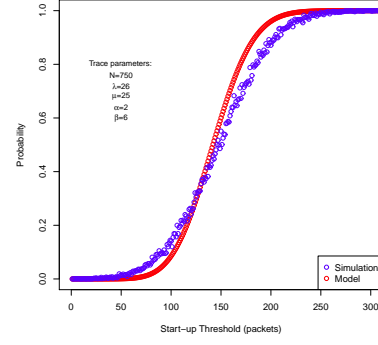


Fig. 6. The probability of no starvation versus the start-up threshold x

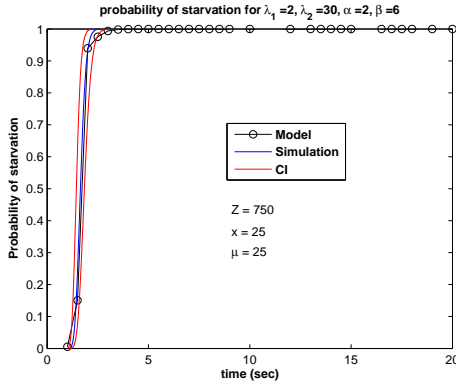


Fig. 4. The probability of starvation for two states

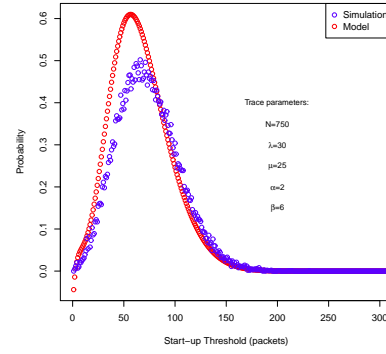


Fig. 7. The probability of having one starvation

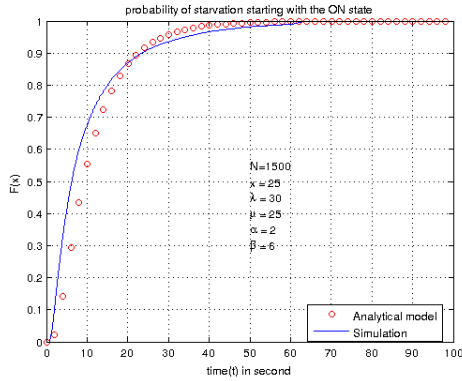


Fig. 5. The probability of starvation for $\lambda > \mu$

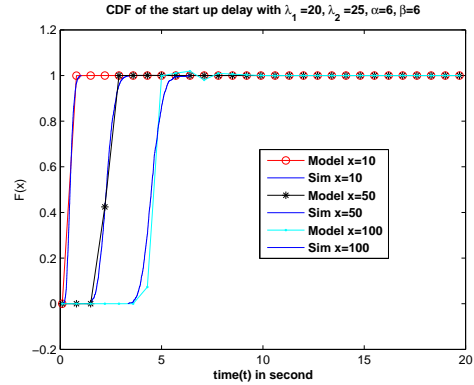


Fig. 8. The CDF of the start up delay for the two states mmpp

figure) no starvation will happen until the end of the video. Since, the curve grows sharply, it is clear that a slight increase in x can greatly improve the starvation probability. In figure 9, we plot the probability of having no more than two starvations with $\lambda_1 = 30$, $\lambda_2 = 0$, $\mu = 25$ irrespective to x . When x is large enough, no starvation will happen until the end of the video session. On the other side, figure 10 shows that the starvation happens for sure when the file size approaches

infinity. The curves of the probability of having one starvation or two starvations increase first, and then decrease to zero. This means that the starvation can be avoided when x is large enough. The two curves have a maximum value at a given start-up threshold or a given file size. So one can choose the threshold x to have exactly one or two starvations. This is an important measure because it allows one to achieve the buffer requirements in setting up the desired values. Indeed,

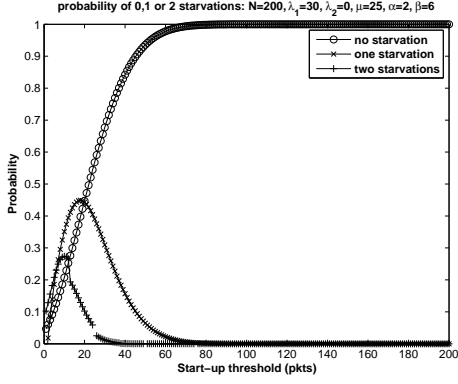


Fig. 9. Probability of 0,1 and 2 starvations versus the start-up threshold x

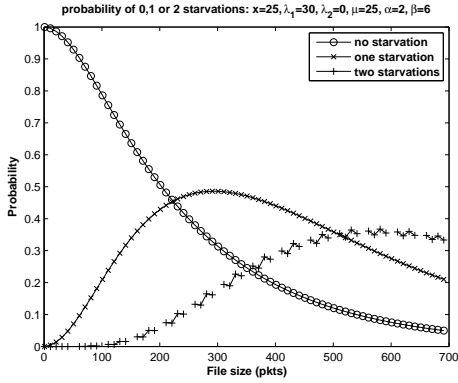


Fig. 10. The probability of 0,1 and 2 starvations versus the file size N

a very small threshold do not help to reduce the starvation probability and very large thresholds do not further reduce the starvation probability. Hence the analytical model aims to predict the player buffer behavior in video streaming sessions. The network parameters and the video size are the framework inputs that can be used to improve the QoE related to user preferences.

C. Optimization of the QoE

In this section, we introduce an optimization problem of the QoE by including different metrics and incorporating user preferences by associated a weight to each metric. We denote by $C(x, Z)$ the cost of a user watching the media stream,

$$C(x, Z) = c_1 \cdot P_{ns}(Z) + c_2 \cdot T_x + c_3 \cdot \Delta F \cdot T$$

where $P_{ns}(Z)$ is the number of starvation, T_x is the start-up delay, ΔF is the lost on video quality and T is the fraction of the total session time spent in low bit-rate. c_1 , c_2 and c_3 are depending on the user preferences of the three metrics (starvation, start-up delay and video quality). Based on the user preferences, we compare the cost of QoE for two scenarios. In scenario 1, the adaptive bitrate streaming is not used and for the second scenario, the adaptive bitrate streaming is used in order to adjust the quality of a video stream

according the available bandwidth. We consider a network that throughput varies between 200Kbps and 400Kbps. For the adaptive bitrate streaming, we have two coding rates depending on the throughput. This leads to two different frame sizes (10kbits, 20kbits). Then we compute the cost $c(x, Z)$ for $x = 20$ and $Z = 1000$.

In fig. 11 and 12, we compare the cost for the two scenarios. For short video duration, the adaptive bitrate streaming is not benefit because there is a less number of starvation and the quality of the video is degraded. But, for the long video duration, the adaptive bitrate streaming becomes interesting because the low coding rate decreases the number of the starvation. In fig. 11 and 12, we can see that the value of the parameter c_3 changes the preference of the user for the quality of the video. For $c_3 = 1$, $c_3 = 1.5$, we use the adaptive bitrate streaming when the size of the file is more than 400, 600 frames respectively. Otherwise, the adaptive bitrate streaming is not necessary.

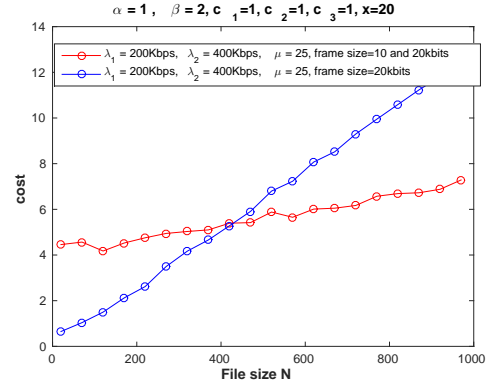


Fig. 11. The cost for progressive streaming versus the adaptive streaming

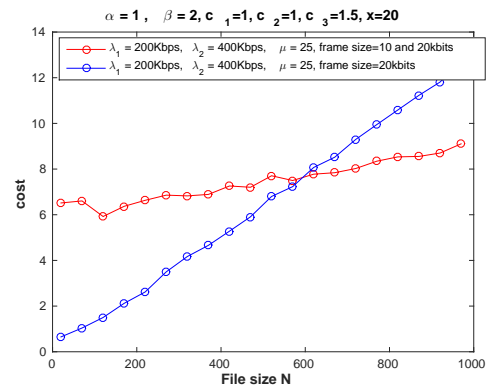


Fig. 12. The cost for progressive streaming versus the adaptive streaming

VIII. CONCLUSION

In this paper, we have proposed a new analytical framework to compute the QoE of video streaming in the network modeled by the Markov Modulated Fluid Model. We found

the probability of starvation and the start-up delay in solving Partial Differential Equations through the Laplace Transform method. This allowed us to compute the number of starvation during the video session that is an important metric of the quality of experience of the user. In addition, we have presented simulation results using ns3 to show the correctness of our model. We have proposed a method to optimize the quality of experience given a trade-off between the player starvation and the quality of the video. These results show that the adaptive bitrate streaming could impact negatively on the quality of the short video duration.

REFERENCES

- [1] Y. Xu, E. Altman, R. El-Azouzi, S. Elayoubi and M. Haddad, *QoE Analysis of Media Streaming in Wireless Data Networks*, Networking Springer Berlin Heidelberg, 2012.
- [2] Y.D. Xu, Y. Zhou and D-M. Chiu, *Analytical QoE Models for Bit-Rate Switching in Dynamic Adaptive Streaming Systems*, IEEE Transactions on Mobile Computing, 2013.
- [3] Y.D. Xu, S.E. Elayoubi, E. Altman, R. Elazouzi and Y. Yu *Flow level QoE of Media Streaming in Wireless Networks*, eprint arXiv:1406.1255, 06/2014.
- [4] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica and H. Zhang, *Developing a Predictive Model of Quality of Experience for Internet Video*, Sigcomm 2013, Hong Kong.
- [5] Y.D. Xu, S.E. Elayoubi, E. Altman and R. Elazouzi, *Impact of Flow-level Dynamics on QoE of Video Streaming in Wireless Networks*, IEEE Infocom, 2013.
- [6] A. Narayanan, and V.G. Kulkarni, *First Passage Times in Fluid Models with an Application to Two Priority Fluid Systems*, Computer Performance and Dependability Symposium, Proceedings IEEE International, 1996.
- [7] F. Dobrian, A. Awan, I. Stoica, V. Sekar, A. Ganjam, D. Joseph, J. Zhan, and H. Zhang, *Understanding the Impact of Video Quality on User Engagement*, Proc. of ACM Sigcomm 2011, Vol.41, No.4, pp. 362-373.
- [8] J. Abate, G.L. Choudhury and W. Whitt, *Computational Probability*, W. Grassman (ed.) Kluwer, Boston 1999.
- [9] Y.D. Xu, E. Altman, R. El-Azouzi, M. Haddad, S. Elayoubi and T. Jimenez, *Analysis of Buffer Starvation with Application to Objective QoE Optimization of Streaming Services*, IEEE Infocom, 2012.
- [10] S. Shah-Heydari, T. Le-Ngoc, *MMPP models for multimedia traffic*, Telecommunication Systems 15, March 1999.
- [11] A. Balachandran, V. Sekar, A. Akella, S. Seshan, I. Stoica, and H. Zhang. A quest for an internet video quality-of-experience metric. In 11th ACM Workshop on Hot Topics in Networks (HotNets-IX), 2012.
- [12] S. S. Krishnan and R. K. Sitaraman. Video stream quality impacts viewer behavior: Inferring causality using quasi-experimental designs. In ACM IMC, 2012.
- [13] H. Hu, X. Zhu, Y. Wang, R. Pan, J. Zhu, and F. Bonomi, *QoE-based multi-stream scalable video adaptation over wireless networks with proxy*, in ICC, 2012
- [14] D. Bethanabhotla, G. Caire, and M. J. Neely, Joint transmission scheduling and congestion control for adaptive video streaming in smallcell networks, arXiv:1304.8083 [cs.NI], 2013.
- [15] V. Joseph, S. Borst, and M. Reiman, Optimal rate allocation for adaptive wireless video streaming in networks with user dynamics, Submitted 2013. Vinay Joseph and Gustavo de Vecian, NOVA: QoE-driven Optimization of DASH-based Video Delivery in Networks, IEEE INFOCOM, April 2014.
- [16] R. Mok, E. Chan, and R. Chang, Measuring the quality of experience of HTTP video streaming, IFIP/IEEE International Symposium on Integrated Network Management, 2011.
- [17] T.H. Luan, L.X. Cai and X. Shen, *Impact of Network Dynamics on User's Video Quality: Analytical Framework and QoS Provision*, IEEE Transactions on Multimedia, Vol. 12, January 2010.
- [18] <http://arxiv.org/abs/1412.2600>

APPENDIX

Proof: When $\lambda_1 \neq \mu$, $\lambda_2 \neq \mu$,

$$Q + s(\omega)R - \omega I = \begin{bmatrix} -\beta + (\lambda_2 - \mu)s(\omega) - \omega & \beta \\ \alpha & -\alpha + (\lambda_1 - \mu)s(\omega) - \omega \end{bmatrix}$$

$(Q + s(\omega)R - \omega I)\phi(\omega) = 0$ and ϕ^k is the eigenvector corresponding to $s_k(\omega)$ according to section IV-A, then

$$\phi^k = \begin{bmatrix} 1 & \frac{\beta + \omega - (\lambda_2 - \mu)s_k(\omega)}{\beta} \end{bmatrix}^T, k = 0, 1.$$

When $\lambda_1 < \mu$, $\lambda_2 \geq \mu$, we use only $\phi^1(\omega)$ in computing the distribution, since $Re(s_0(\omega)) > 0$. Thus the distribution $\tilde{H}_1(x, \omega)$ becomes

$$\begin{bmatrix} \tilde{H}_{11}(x, \omega) \\ \tilde{H}_{21}(x, \omega) \end{bmatrix} = a_{11}e^{s_1(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{\beta} \end{bmatrix}$$

$a_{11} = \tilde{H}_{11}(0, \omega)$ and $\tilde{H}_{11}(0, \omega) = 1$ because if we start without packets in the buffer in state 1, we'll have starvation with probability 1 within the same state. So $a_{11} = 1$, that yields the result of the theorem. The same proof holds in the case $\lambda_2 < \mu$, $\lambda_1 \geq \mu$ by interchanging λ_1 and λ_2 .

When $\lambda_1 < \mu$, $\lambda_2 < \mu$, we use both $\phi^0(\omega)$ and $\phi^1(\omega)$. Thus we have

$$\begin{aligned} \tilde{H}_1(x, \omega) &= \begin{bmatrix} \tilde{H}_{11}(x, \omega) \\ \tilde{H}_{21}(x, \omega) \end{bmatrix} \\ &= a_{11}e^{s_0(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_0(\omega)}{\beta} \end{bmatrix} + \\ &\quad a_{21}e^{s_1(\omega)x} \begin{bmatrix} 1 \\ \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{\beta} \end{bmatrix} \end{aligned}$$

Using the initial condition $\tilde{H}_{11}(0, \omega) = 1$ and $\tilde{H}_{21}(0, \omega) = 0$, we solve the following system

$$\begin{cases} a_{11} + a_{21} = 1 \\ a_{11} \frac{\beta + \omega - (\lambda_2 - \mu)s_0(\omega)}{\beta} + a_{21} \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{\beta} = 0 \end{cases}$$

and get

$$\begin{aligned} a_{11} &= \frac{\beta + \omega - (\lambda_2 - \mu)s_1(\omega)}{(\lambda_2 - \mu)(s_0(\omega) - s_1(\omega))} \\ a_{21} &= \frac{\beta + \omega - (\lambda_2 - \mu)s_0(\omega)}{(\lambda_2 - \mu)(s_1(\omega) - s_0(\omega))} \end{aligned}$$

The same proof yields for $\tilde{H}_2(x, \omega)$. ■

